Limits on Threshold and "Sommerfeld" Enhancements in Dark Matter Annihilation

Mihailo Backović and John P. Ralston

Department of Physics & Astronomy,
The University of Kansas, Lawrence, KS 66045

We find model-independent upper limits on rates of dark matter annihilation in galactic halos. The Born approximation generally fails, while exotic threshold enhancements akin to "Sommerfeld factors" also turn out to be baseless. The most efficient annihilation mechanism involves perturbatively small decay widths that have largely been ignored. Widths that are very small compared to TeV mass scales suffice to cause large enhancements in the velocity averaged cross sections. Bound state formation in weakly coupled theories produces small effects due to wave function normalizations. Unitarity shows the Sommerfeld factor cannot produce large enhancements of cross sections, and serves to identify where those approximations break down.

I. THRESHOLD ENHANCEMENTS

There is great interest in recent data from the PAMELA [1], FERMI [2] and PPB-BETS [3] experiments. The observations suggest a significant signal in excess positron production in galactic halos, as long suggested by the HEAT [4] and ATIC [5] experiments. Possible explanations range from exotic mechanisms [6], uncertain features of pulsars [7], to dark matter decays [8] and dark matter annihilation [9].

In considering annihilation there are puzzles from comparing predictions of relic densities with rates of particle production in the current era. This has led to invoking more or less exotic threshold enhancements under the catch-phrase of "Sommerfeld factors" [10, 11].

One reason to appeal to a Sommerfeld factor is to boost cross sections of TeV-scale particles well above Born-level estimates. We find the starting point of Born-level cross sections is not a good approximation for much different reasons. Basic facts of finite width particle physics substantially revise estimates of annihilation rates in galactic halos. We find that annihilation of TeV-scale dark matter with typical electroweak couplings can actually saturate unitarity limits over the observable range. We also obtain upper limits to halo annihilation rates that do not depend on fine details of the dark matter velocity distribution.

Non-relativistic scattering amplitudes can be classified by their analytic properties in the complex momentum plane. Stable bound states are described by poles on the positive imaginary axis. It follows that stable bound states produce no remarkable enhancement of annihilation rates in the physical region of real momentum k. *Metastable particles* or *resonances*, described by poles of finite width, are in no way comparable with stable bound states, because everything observable (and potentially large) is a strong function of the width.

Unless one is considering an absolutely stable intermediate state, *all intermediate states in particle physics have a finite lifetime.* Pursuing the consequences of finite lifetimes with galactic halo kinematics re-directs attention from exotic mechanisms to ordinary physics. There are two salient cases. If the width of an intermediate annihilation state is limited by the initial state velocity v, then the peak of the cross section goes like $1/v^2$. This case produces the largest reaction rates in halos, and most conservative bounds. If the width of the intermediate state is constant, the peak of the cross section goes like $1/v$. In these and intermediate cases the peak cross section actually dominates the entire halo velocity distribution for a surprisingly broad range of dark matter parameters. As a result, our more conservative bounds merge smoothly with reasonable estimates predicting surprisingly large rates.

II. BREIT-WIGNER FORMULAS

Relic particles trapped in galactic halos will be non-relativistic, with velocities $v \sim 10^{-3}$. There are several distinctly different non-relativistic "Breit-Wigner" formulas. Most Breit-Wigner cross sections σ_{res} can be cast into the form

$$\sigma_{res} = \frac{4\pi v^N}{k^2} \frac{(\Gamma/2)^2 B_i B_f}{(E - E_{res})^2 + (\Gamma/2)^2} = \frac{4\pi v^N}{k^2} BW(\Gamma, E_{res}).$$
(1)

Here B_i and B_f are the branching fractions to the initial and final state, and k is the momentum of an initial state particle in the center of mass frame. Different values of the parameter $N = 0, 1$ distinguish two classic limits:

Phase Space Limited Case, $N = 0$: It is common for $2 \rightarrow 2$ non-relativistic physics to be quasi-elastic. In particular, the *final* state phase space may be severely limited by the *initial* state velocity v. Ignoring spin and matrix elements, the Lorentz-invariant phase space integral $LIPS$ for two particles of momentum p_f, p_f and

mass m_f: is

$$LIPS = \int \frac{d^3p_f}{2p_f^0} \frac{d^3p_f}{2p_f^0} \delta^4(Q - p_f - p_f)$$
$$= 2\pi\sqrt{1 - 4m_f^2/Q^2} = 2\pi v_f. \qquad (2)$$

Here v_f is the final state velocity of either particle in the CM frame. When initial and final state masses are comparable, and the 2-body states dominate, the total width $\Gamma \sim \kappa v_f \sim \kappa v$, where κ absorbs coupling constants and matrix elements. Incorporating the explicit velocity dependence with an s-channel propagator leads to Eq. 1 with $N = 0$. Note that the peak of the cross section scales like $1/(m^2 v^2)$, making this case potentially capable of saturating elastic unitarity bounds.

Relativistic Phase Space Case, $N = 1$: Anihilation may also proceed to final states which are ultra-relativistic. Then the square root in Eq. 2 approaches 1, and the partial width $\Gamma_f \sim constant$ in this limit. Any other kinematic situation where Q^2/m_f^2 goes to a finite constant as $v \to 0$ will produce the same outcome. This includes the "exoergic" resonances long known in low-energy nuclear physics, and associated with the "$1/v$ law" of low energy cross sections. These cross sections do not increase as fast as unitarity would allow as $v \to 0$.

The difference between $1/v$ and $1/v^2$ velocity dependence is dramatic. Yet is only part of the story, because resonances may produce large cross sections either way. For example, neutron absorption cross sections on Gadolinium-157 exceeding *one hundred million barns* have been observed. [12]. This comes in the seemingly mild $1/v$ case not impinging on a unitarity limit. The experimental stunt simply exploits neutrons with grossly small velocities of order 3 meters per second. In much the same way, galactic halo velocities of order 10^{-3} are grossly small on the scale of particle physics. The combination of low speed halo kinematics and very ordinary widths produces surprisingly large enhancements.

III. GENERAL LIMITS ON THE VELOCITY AVERAGED BREIT-WIGNER CROSS SECTION

The halo annihilation rate via a single $s-$wave resonance is governed by the velocity-weighted cross section $\langle \sigma v \rangle_{res}$:

$$\langle \sigma v \rangle_{res} = \int dv\, v \frac{4\pi v^N}{m_X^2 v^2}$$
$$\times \frac{(\Gamma/2)^2 B_i B_f}{(m_X v^2/2 - m_X v_{res}^2/2)^2 + (\Gamma/2)^2} \Phi_{halo}(v).$$

Here $\Phi_{halo}(v) = dN/dv^3$ is the normalized dark matter relative velocity distribution, assumed from astrophysics to be a smooth function on the scale of 100-500 km/s. In an isothermal halo model the velocity distribution is in equilibrium,

$$\frac{dN}{d^3k d^3x} = \frac{constant}{E_0} e^{-E/2E_0};$$
$$\frac{dN}{dv} = 4\pi \frac{v^2}{(2\pi v_0^2)^{3/2}} e^{-v^2/2v_0^2}. \qquad (3)$$

While the actual velocity distribution is uncertain the phase space factors of v^2 are general. The isothermal halo will illustrate the method, but none of our upper bounds depend on it.

The rate $\langle \sigma_{res} v \rangle_{res}$ is a function of E_0, E_{res}, Γ and m_X. If other scales are expressed in units of $m_X \sim TeV$ the conjunction of several rapidly varying functions makes analysis troublesome, as noted by Griest and Seckel [13]. However in the present universe the halo energy $m_X v_0^2/2 \sim 10^{-6} m_X$ is rather small on particle physics scales. It is natural to rescale variables in units of the halo characteristic energy, defining

$$\gamma_0 = \frac{\Gamma}{2E_0}; \quad \delta_0 = \frac{E_{res}}{E_0}.$$

Assuming the equilibrium distribution, some algebra gives

$$\langle \sigma v \rangle_{res} = \frac{2^{2-N}(2\pi)^{\frac{N+1}{2}} v_0^{N-1}}{m_X^2} I_N(\gamma_0, \delta_0), \qquad (4)$$

where

$$I_N(\gamma_0, \delta_0) = \frac{1}{2^{1-N}(2\pi)^{N/2}} \int_0^\infty dz\, z^{\frac{N}{2}} \frac{\gamma_0^2 e^{-z/2}}{(z - \delta_0)^2 + \gamma_0^2}. \qquad (5)$$

Note that $I_N(\gamma_0, \delta_0)$ is analytic for all $\gamma_0 > 0$ and δ_0 regardless of the sign of δ_0. It can be computed exactly in terms of Exponential Integral (Ei) functions. We found it more useful to observe that $I_N(\gamma_0, \delta_0)$ has certain *absolute upper limits* for all possible values of $\gamma_0 > 0$ and δ. Consider the derivative $\partial I_0(\gamma_0, \delta_0)/\partial\gamma_0$:

$$\frac{\partial I_0(\gamma_0, \delta_0)}{\partial\gamma_0} = \int dz\, \frac{2\gamma_0(\delta_0 - z)^2}{(\gamma_0^2 + (\delta_0 - z)^2)^2} e^{-z}. \qquad (6)$$

Since the integrand above is positive definite, the integral achieves its maximum at $\gamma_0 \to \infty$. For $\gamma_0 >> 1$ the integration becomes trivial, yielding $I_N(\gamma_0, \delta_0) \leq 1$. A stronger limit notes the integrand of Eq. 5 is cut off for $z \lesssim \gamma_0$ when $\gamma_0 \lesssim 1$, $\delta_0 \lesssim 1$, implying $I_N(\gamma_0, \delta_0) \lesssim 1 - e^{-C\gamma_0}$, where C is a constant. Numerical work shows that for all parameters

$$I_0(\gamma_0, \delta_0) \leq 1 - e^{-\frac{\pi}{2}\gamma_0}, \qquad (7)$$
$$I_1(\gamma_0, \delta_0) \leq 1 - e^{-\frac{\pi}{4}\gamma_0}.$$

These are close to equality for positive $\delta_0 << 1$. Figure 1 shows a plot of $I_N(\gamma_0, \delta_0)$ for a wide range of γ_0, δ_0 and how the integral approaches the upper bound.
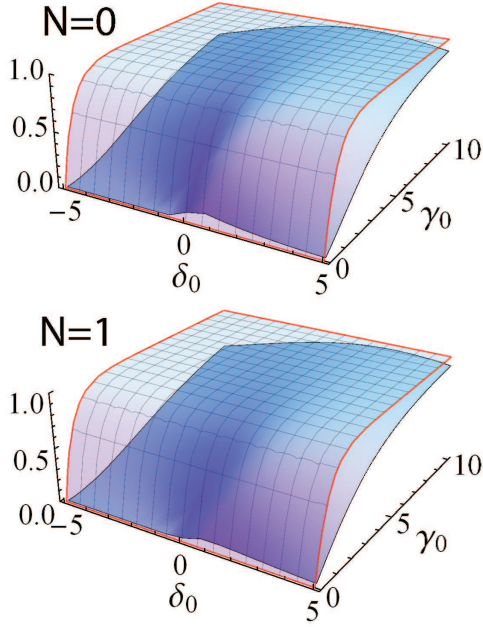
FIG. 1: The integral $I_N(\gamma_0, \delta_0)$ (dark shaded) and upper limits cited in the text (transparent mesh). By Eq. 4 the rate constant is related via $\langle \sigma v \rangle_{res} \sim v_0^{N-1} I_N / m_X^2$.

The positivity property of Eq. 6 holds for all halo distributions. The upper limit $BW \to 1$ produces a universal inequality:

$$\langle \sigma v \rangle_{res} < \frac{4\pi \langle 1/v^{1-N} \rangle}{m_X^2}. \qquad (8)$$

The expected value $\langle 1/v \rangle$ is relative to the distribution $\Phi_{halo}(v)$, not dN/dv. If the equilibrium distribution is assumed, then

$$\langle \sigma v \rangle_{res} < \frac{2^{2-N}(2\pi)^{\frac{N+1}{2}} v_0^{N-1}}{m_X^2}(1 - e^{-\pi\gamma_0/2^{N+1}}) \qquad (9)$$

The result is a possible significant enhancement factor (EF) ("boost factor") for annihilation rates. The enhancement factor is defined relative to a typical Born approximation $\sigma_{Born} = 4\pi\alpha_X^2/m_X^2 v^{2-N}$:

$$EF = \frac{\langle \sigma v \rangle_{res}}{\langle \sigma v \rangle_{Born}} \lesssim \frac{1}{\alpha_X^2}. \qquad (10)$$

Note that the upper limit does not depend on the position of the resonance nor on any halo properties.

1. N = 0 Enhancement Factors

For $N = 0$, Eq. 10 leads to substantial enhancements approaching the unitarity bound when the fundamental width Γ is large enough. Obtaining a "large
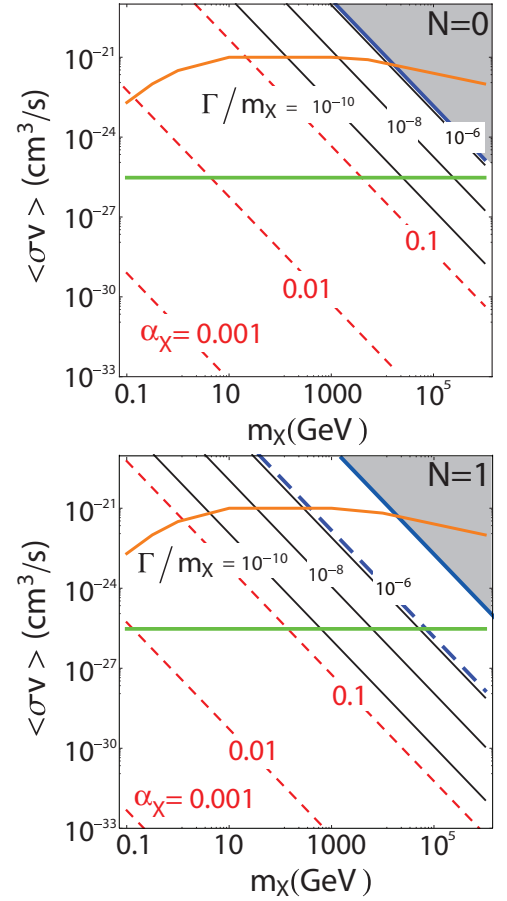

FIG. 2: Upper limits (diagonal lines) of resonantly enhanced annihilation rate $\langle \sigma v \rangle_{res}$ in the isothermal halo distribution. Solid curves (black online) are computed with fixed Γ/m_X. Gray triangle in upper right is the unitarity bound. The thick dashed curve (blue online) is the maximum value for the cross section for $N = 1$. Thin dashed curves (red online) show $\langle \sigma v \rangle_{res}$ computed for bound state processes using $\Gamma = \alpha_X^5 m_X/2$ and $E_{res} = -m_X \alpha_X^2/4$. Middle curve (orange online) is the neutrino-based upper limit of Ref. [17]. Horizontal line (green online) is a conventional lower bound $\langle \sigma v \rangle \sim 3 \times 10^{-26} cm^3/s$.

enough" width from a weakly coupled theory might appear special. Yet remember that halo annihilations are driven by the width in units of the rather small scale $E_0 \sim 10^{-6} m_X$. For TeV-scale dark matter a width $\Gamma \gtrsim$ MeV is large enough to dominate the halo width and make $BW(\Gamma, E_{res}) \sim 1$. Recall that the J/ψ has a width of order 0.1 MeV and is exceedingly "narrow". For an elementary particle on any mass scale of GeV-TeV *not to have* widths exceeding $10^{-6} m_X$ requires special conspiracies or selection rules.

Figure 2 shows that even a tiny value of $\Gamma/m_X \sim 10^{-8}$ can produce rates much larger than the oft-cited value $\langle \sigma v \rangle \sim 3 \times 10^{-26} cm^3/s$. It is a new insight that merely including physics of widths tends to saturate unitarity bounds in halo annihilation.
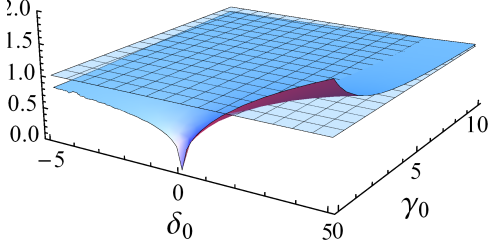
FIG. 3: Ratio of I_1/I_0 (shaded area) compared to the uniform value of 1 (mesh)

2. $N = 1$ Enhancement Factors

Equation 8 highlights a factor of $\langle 1/v \rangle$ absent with a relativistic phase space ($N = 1$). To a first approximation the ratio of the $N = 1$ case relative to the $N = 0$ case is of $O(v_0)$. This is made more precise using Figure 3, which shows a plot of the calculated ratio of integrals I_1/I_0 that remains. This ratio is of order unity for most of the parameter space, except the regions where $\gamma_0 << 1$. Once again, only when widths are very tiny do resonance widths *not* tend to swamp the halo distribution.

While representing stronger limits, the bottom panel of Figure 2 again shows significant enhancements over a broad range of parameters of current interest. The difference between I_1 and I_0 tends to disappear whenever Γ/m_X is not exceptionally small. In the next Section we turn to the metastable bound state case, which does happen to exhibit exceptionally small widths on general grounds.

A. Metastable Bound States, and Narrow Resonances

The case of annihilation passing through intermediate metastable bound states has generated great interest. This case is different and deserves a separate discussion. Suppose dark matter interacts with a light messenger particle of mass μ, with coupling-squared α_X. If the interaction is attractive, which is readily arranged for particular spins, then non-relativistic physics predicts *there is always a bound state* for sufficiently large coupling. The conditions are

$$\alpha_X \gtrsim \kappa \frac{\mu}{m_X},$$

where κ is a constant of order one. The demonstration is an easy variational calculation using a ground state Hydrogenic wave function. A helpful discussion is also given in Ref. [14]. For parameters $m_X \sim TeV$ and $\mu \sim GeV$ bound state formation needs $\alpha_X \gtrsim 10^{-3}$, which is well within the electroweak-scale couplings of most models.

Yet just as above, everything about any significant enhancement depends strongly on the *width*, and won't proceed without it. To estimate widths, first note that bound states are spatially large for small coupling constant α_X. The size of a weakly coupled bound state is roughly estimated by the "Bohr radius" a_0, where

$$a_0 \sim 1/m_X \alpha_X.$$

Similarly, the binding energy is $E_{res} \sim m_X \alpha_X^2$. Next recall that the Schroedinger wave function at the origin $\psi(0)$ determines the width via $\Gamma \sim |\psi(0)|^2 \sigma_c$ where σ_c is a continuum cross section.

The wave function at the origin is set by the inverse of the size of the bound state:

$$|\psi(0)|^2 \sim a_0^{-3} \sim \alpha_X^3.$$

The continuum annihilation cross section $\sigma_c \sim \alpha_X^{1+A}$ for $A > 0$ depending on the model. For reference the annihilation rates of ortho (para) positronium via three (two) photons go like $\alpha_{em}^6 (\alpha_{em}^5)$. Thus bound state widths follow a general pattern

$$\Gamma \sim \alpha_X^{4+A} m_X \lesssim 10^{-8} m_X.$$

The right hand side is a fair upper limit for $\alpha_X \sim 10^{-2}$. Restricted phase space factors and branching ratios can only reduce this. Comparing $E_0 \sim 10^{-6} m_X$, we find that $\gamma_0 << 1$ is by far the generic case for annihilation from a bound state. As a consistency check, consider the definite case of spin-1/2 dark matter interacting with vector particles. Nature has already done this calculation with the J/ψ decay via gluons, which has $\Gamma_{J/\psi}/m_{J/\psi} \lesssim 10^{-4}$. The J/ψ is sufficiently heavy that the perturbative phase space factors are driven by dimensional analysis, as expected for TeV-scale physics. The raw J/ψ ratio needs to be re-scaled by $(\alpha_X/\alpha_s)^{4+A} \sim 10^{-4}$, which gives satisfactory agreement.

When $\gamma_0 << 1$ it is a good approximation to replace $BW(\Gamma, E_{res}) \sim \pi(\Gamma/2)\delta(E - E_{res})$. A short calculation then gives

$$EF(\gamma_0 << 1) = \frac{\pi \Gamma/2}{\alpha_X^2 m_X v_{res}^{2-N}} \frac{\Phi_{halo}(v_{res})}{\langle \frac{1}{v^{1-N}} \rangle}, \quad (11)$$

where $E_{res} = m_X v_{res}^2/2$. This formula has no singularity as $v_{res} \to 0$ because $\Phi_{halo}(v_{res}) \sim v_{res}^2$ has compensating factors from phase space (Eq. 3). If a metastable bound state resonance lies *above threshold* in an expected electroweak range the effects are quite small. Taking $E_{res} = m_X \alpha_X^2/2 \sim 10^{-4} m_X$, and the equilibrium halo model with scale $v_0 = 10^{-3}$, the factor $e^{-E_{res}/E_0} \sim e^{-100}$ is too small to consider further. When the resonance is *below threshold* it must have width $\Gamma \gtrsim |E_{res}|$ to intrude into the physical region. Since Γ is proportional to several powers of α_X compared to E_{res} this case can also be set aside. If $E_{res} \to 0$ with $\Gamma >> E_{res}$ is contemplated, it implies the decay time scale is much less than a binding

(orbital) time scale, which is not consistent with bound states forming in the first place.

An exponentially small suppression can be avoided by adjusting the binding into the range probed by the halo velocity. For example choose $\alpha_X \sim 10^{-3}$. This device rapidly loses consistency because the bound state criterion $\alpha_X \gtrsim \mu/m_X$ needs couplings not too small. If a bound state is tuned to the vicinity of the peak, then the halo factors will be order unity. Meanwhile there remains in Eq. 11 an overall factor of $\Gamma/(m_X \alpha_X^2) << 1$. Figure 2 compares the upper limits from annihilation of continuum processes ($\gamma_0 \gtrsim 1$ generically) to processes proceeding via the bound state ($\gamma_0 \lesssim 10^6 \alpha_X^5$) using the isothermal halo and conservative values $B_i B_f \to 1$. Viable enhancement mechanisms should also respect the neutrino-based bounds of Mack, Beacom and Bell [17] included in the Figure. In case of $N = 0$, a bound state could cause large cross section enhancements, but only for couplings $\alpha_X \geq 0.1$ which are beyond the stable perturbative regime. In case of $N = 1$ the limits for bound states are even tighter.

Figure 2 shows that a single bound state with perturbative couplings has no chance of causing significant enhancements. Except for strong coupling, there is *no dynamical mechanism to generate large enhancement factors from non-relativistic bound state resonances in the current universe.* The conclusion does not depend on the spin or quantum numbers of new physics, and is too strong to escape by adding up several resonances, unless they are so numerous their numbers alone overcome small couplings, as for KK modes.

1. Breit Wigner Effects on Relic Abundance

Relic abundance is a different topic than halo annihilation. Ibe, Murayama, and Yanagida [15],and Guo and Wu [16] have calculated thermal evolution for the case of a narrow state close to threshold. Their model cross section is essentially equivalent to our $N = 1$. The resonance position is close enough to threshold for its effects to overlap into the physical region during relic evolution. They find that even a tiny ratio of width to resonance invariant mass, denoted by $\gamma = \Gamma/M_{res} \lesssim 10^{-3}$, produces significant effects on relic densities compared to traditional constant cross section approximations. The sense of this effect causes a relative decrease in annihilation rates in the early universe, which tends to leave too much relic. To keep the relic density Ω_X fixed, avoiding over-closure, Refs. [15, 16] introduce "boost factors" to correct the normalization parameters of the cross section $\sigma_0 \sim B_i B_f/m_X^2$. When those boost factors are applied directly to halo annihilation, they develop much larger cross sections than the well-known cosmological value $\sigma_0 \sim 10^{-9} GeV^{-2}$, which might be relevant to the PAMELA-ATIC observations.

However, in general it is not possible to go from the relic calculation to the halo calculation directly in this manner. The halo annihilation rate $< \sigma v >$ has a new and separate sensitivity that is *a priori* disconnected from relic calculations. The halo estimates are driven by the new parameter $\gamma_0 = \gamma M_{res}/2E_0$. As long as $\gamma_0 >> 1$, the upper bounds on the halo annihilation cross section will be saturated. This key feature is conceptually absent if the halo annihilation cross section is simply re-scaled by factors invoked for relic evolution. Thus the reported "boost factors" of the relic calculations do not take into account the Breit-Wigner effects on halo annihilation we have found. This explains why the suggestion [15, 16] that very small Γ/M_{res} is necessary or tends to enhance halo rates is not general, and appears different from our conclusion. It is clearly possible to find models and parameter regions where both, or neither of the correct relic density and halo enhancement phenomena can be accommodated.

We note that relic densities are also subject to many uncertainties of galaxy formation and the other boost factors representing "clumpiness". For purposes of confronting experimental data, it seems best to separate the problems of halo annihilation and relic evolution entirely, despite mathematical similarities in how they are calculated.

IV. SOMMERFELD FACTORS

We have shown that Breit-Wigner width effects of typical particle physics type can be surprisingly large, while bound state effects have little chance to compete. *Sommerfeld factors* have also been claimed as a mechanism to produce large enhancements not involving particle widths [11].

Given an $s-$wave cross section σ^0, which has been computed in the plane wave basis, the Sommerfeld-based recipe to include Coulomb wave effects is to make a replacement

$$\sigma^0 \to \sigma^0 S(v, \alpha);$$
$$S(v, \alpha) = \frac{\alpha}{v} \frac{2\pi}{1 - e^{-2\pi\alpha/v}}.$$

Here α is the fine structure constant. Since cross sections contain many other terms of different orders in α^j/v^k, together with logarithmic type dependence, the recipe is an approximation by *re-summation* of selected contributions [18].

2. Motivation for Re-summation

Non-relativistic QED has complicated logarithmic and power-behaved infrared singularities. Singular terms must be summed or controlled in some way to avoid upsetting perturbation theory. There are reasons to believe that the leading singularities appear [28] order by order as a series in α^j/v^j. Evidently such a series is summed in $S(v, \alpha) = S(\alpha/v)$.

Sub-leading terms are dropped in any re-summation - for example, a term of order α^j/v^k with $j > k$ is sub-leading as $v \to 0$. The fact that infinitely many sub-leading terms exist comes from the fact that it is always possible to add a photon exchange loop to any diagram, and all loops have some integration region *not* singular as $v \to 0$.

The purpose of re-summation is to extend the reach of perturbation theory into the difficult, non-perturbative regime. Some examples illustrate typical limitations. Expand $S(v, \alpha) \sim 1 + \pi(\alpha/v) + \pi^2\alpha^2/3v^2 + \dots$. For $\alpha = 10^{-2}$, and $v = 5 \times 10^{-2}$ the third term is $\pi^2/75 \sim 0.13$. It happens to be larger than a typical non-singular term of order α; retaining it is well-motivated for these kinematics. The series is also stable in this regime: a high order term $(\pi\alpha/v)^{10} \sim 0.0096$ is small. Yet at the smaller velocity $v = 5/1000$, the sub-leading correction $\alpha^{11}(\pi/v)^{10} \to 958956.0$ is not included, is hardly small, and the leading-order re-summation fails. It is the nature of such re-summation that self-consistency breaks down in the region $\alpha/v \gtrsim 1$, exactly where $S(v, \alpha) \gtrsim 1$.

If one believes a re-summation recipe might generate corrections of order 30-50%, say, there's seldom any reason to invoke it for factors of "10" or more. Yet recent treatments of dark matter annihilation have imagined the Sommerfeld factor to be very general. It has been held responsible for extremely large enhancement factors of $S >> 10$, while also coming from any generic interaction involving light Yukawa particles [20]. The perception comes from a practice of citing continuum Coulomb normalization factors $|\psi_C|^2(0)$. Since Coulomb normalizations are known exactly, the procedure has been thought to be "exact."

We have traced early literature to find several logical and historical contradictions. Guth and Mullin highlighted the approximations in 1951 [21]. *In lowest order approximation*, but while using Coulomb wave functions $\psi_{C\,i}$ for a basis, one encounters matrix elements M of the form

$$M = \int d^3x\, \psi_{C\,2}^* V \psi_{C\,1}.$$

Insert complete sets of momentum eigenstates $|k\rangle$:

$$M = \int d^3k\, d^3k'\, \psi_{C\,2}^*(k) V_{kk'} \psi_{C\,1}(k').$$

The Coulomb wave functions are sharply peaked in the vicinity of certain momenta k_1, k_2, identified by taking the limit $\alpha = 0$. Assume the plane-wave matrix elements $V_{k_1,\,k_2} = \langle k_1|V|k_2\rangle$ are relatively smooth functions of momentum transfer. Make a rough approximation moving $V_{k_1 k_2'}$ outside the integral:

$$\begin{aligned} M &\to V_{k_1,\,k_2} \int d^3k\, \psi_{C\,2}^*(k) \int d^3k'\, \psi_{C\,1}(k), \\ &\to V_{k_1,\,k_2} \psi_{C\,2}^*(x=0)\psi_{C\,1}(x=0). \end{aligned} \quad (12)$$

In the last line $\psi_C(x = 0)$ appears as the coordinate-space wave function at the origin, "improving" the plane wave calculation. Inserting the analytically known normalization of $|\psi_C|^2$ then produces the factor $S(v, \alpha)$ for the cross section.

The operations of separating the collision and wave function integration into products is one of *leading power factorization*. It is used in QCD calculations separating "hard" and "soft" regions of perturbation theory, but in that case while attempting to be systematic. Careful work with positronium annihilation [22] does not use the factorized approximation. Instead, reference to re-summation is made after the full calculations are carried out. An early work [22] on one-loop corrections to positronium decay states that "Coulomb effects are included by this (factored) method to all orders in e^2, *though only, of course, approximately*." (Italics are ours.)

What did Sommerfeld actually do? We consulted his 1931 article, in German, to see it introduced exact Coulomb wave functions to calculate bremsstrahlung, while it never suggested factorization. It is a *tour de force* of early quantum theory; consulting it for a renormalization factor actually perpetuates a normalization mistake. Cross sections are defined by *ratios relative to a flux computed with a given normalization*. The overall normalization of physical states cancels out in total cross sections: and so Eq. 12 is not only approximate, it is incomplete.

Elwert's 1939 dissertation [23] recognized this, as as did Guth in 1941 [24]. These papers abandoned Sommerfeld's calculation and used the *ratio* of two in- and out- Coulomb factors as an approximate factorized ansatz. "Elwert factors" are used in atomic and molecular physics to cancel spurious pre-factors going like v from other approximations, but only when their effects are not too large. Experimental confrontation of the Elwert factor finds errors of relative order unity in the region the factors are of order unity [23, 25]. Elwert and collaborators find this kind of breakdown reasonable [23]. In no event are very large corrections ever credited.

3. Multiplicative Factors Must Fail

In retrospect, we find the concept behind generating singularities via multiplicative factors questionable on general grounds.

A general scattering amplitude has the partial wave expansion

$$f(\theta, k) = \frac{1}{k} \sum_l (2l + 1) f_l(k) P_l(cos\theta).$$

For each partial wave cross section σ_l of angular momentum l, elastic unitary gives the upper limit

$$\sigma_l = \frac{4\pi(2l+1)|f_l(k)|^2}{k^2} \leq \frac{4\pi(2l+1)}{k^2}.$$

This summarizes the unitarity bound of Ref. [26]. Since each partial wave has a finite cross section, *no partial*

wave can possibly have a singularity. "Improving" the *s*-wave cross section - or any particular partial wave cross section - by singular terms of order α_X/v then contradicts unitarity for $\alpha_X/v \gtrsim 1$. This is just the same region where the claimed Sommerfeld factor $S(v, \alpha) >> 1$.

Can one escape the contradiction by appealing to small α_X? It seems not: No small value for α_X is used in the logic of an exact normalization citing Sommerfeld. Small coupling is also no protection from internal inconsistency. Unitarity and analyticity in perturbation theory are exact facts maintained in a systematic way, order by order, regardless of the size of the coupling constant, small or large. When violated, it shows the calculation was bad, just as indicated by sub-leading terms[29].

This problem of consistency is different from the one previously recognized. Dark matter interactions have finite range, while the infrared singularities of re-summation come from infinite range. To account for this the authors of Ref. [11] argued that for a finite range potential, a Sommerfeld enhancement would saturate when the deBroigle wavelength of colliding particles would be larger than the range of the force. A related statement actually follows from a WKB approximation: when the de Broglie wavelength is tiny compared to the range, *and* one works with wave functions at short distance, the range effects of a Yukawa potential $e^{-\mu r}/r$ drop out. Note that the range criteria do not depend on a coupling constant, and also don't specify any particular angular momentum channels. Yet the singularities of scattering amplitudes, and particularly Coulomb singularities, do depend on the couplings and angular momentum channels. Whatever the scale where analogies between massless and massive models break down, the facts of partial wave unitarity are more general, and take precedence. They preclude large enhancement in any particular channel of fixed angular momentum.

This leads to another useful bound. Replace every partial wave by the one with largest magnitude $|f_{max}|$. Sum them up: The result is

$$\sigma \lesssim \frac{4\pi}{k^2} \sum^{l_{max}} (2l+1)|f_{max}|^2 = \frac{4\pi|f_{max}|^2(l_{max}+1)^2}{k^2} \quad (13)$$

This is a strict upper limit. The Sommerfeld factor is not a resonance, and in the absence of resonances $|f_{max}|^2$ for every partial wave is small in weakly coupled theories, making σ small. The notion of canceling a perturbative factor of α_X^2 with an enhancement of 10^4 (say) needs $l_{max} >> 10^2$ partial waves. This is difficult to conceive - or at least a high burden of proof - for finite-range, perturbatively coupled Yukawa models. It is even more problematic that the angular momentum involved

in annihilation is strictly limited by the quantum numbers of intermediate states. When the intermediate state consists of a single particle, the angular momentum is bounded by its spin, closing the door on large spin sums.

Since each partial wave is finite, how do some Coulomb-dominated processes, such as Rutherford scattering, actually become singular as $v \to 0$? As Wigner [27] and many others have noted, the Coulomb singularity is very special. On semi-classical grounds (actually the facts of Legendre series), in Eq. 13 the upper limit $l_{max} \sim r_{max}k$, where r_{max} is the range of the potential. This gives

$$\sigma \lesssim 4\pi r_{max}^2 |f_{max}|^2 \lesssim 4\pi r_{max}^2.$$

The Coulomb singularity occurs because (1) the effective range $r_{max} \to \infty$, and (2) an infinite number of partial waves actually can contribute. Closely related is Wigner's classic theorem [27] that power-law potentials $V(r) \gtrsim 1/r^2$ are needed to develop any kind of singularity.

V. CONCLUDING REMARKS

We have explored significant effects in halo annihilation rates due to natural widths of intermediate states. The problem is intricate due to subtle interplay of energy scales. The Born approximation almost always fails, giving gross underestimates of reaction rates. Tiny values of galactic halo velocities reverse an assumption that propagator widths might be "small corrections." That perception comes from comparing widths to particle masses, and does not capture the important features of halo annihilation.

Given that TeV-scale particles with typical electroweak type couplings may easily have $\Gamma/E_0 >> 1$, Breit-Wigner factors of ordinary radiative corrections must generally be taken into account. Consistency of rates in particular channels, such as the apparent dominance of leptons, still needs to be considered model by model. The fact that merely including basic physics of widths may be quite significant. It revises the basic picture of annihilation scenarios confronting the PAMELA-FERMI-PPB-BETS data in a positive way that increases the possibilities to find new physics.

[1] O. Adriani *et al.* [PAMELA Collaboration], Nature **458**, 607 (2009) [arXiv:0810.4995 [astro-ph]].

[2] E. A. Baltz *et al.*, JCAP **0**807, 013 (2008) [arXiv:0806.2911 [astro-ph]].

[3] S. Torii *et al.* [PPB-BETS Collaboration], arXiv:0809.0760 [astro-ph].

[4] S. W. Barwick et. al. , [HEAT Collaboration] Astrophys.J. 482 (1997) L191-L194

[5] J. Chang et. al. , Nature 456. (2008)

[6] A. A. Andrianov, D. Espriu, P. Giacconi and R. Soldati, JHEP **0**909, 057 (2009) [arXiv:0907.3709 [hep-ph]]; J. McDonald, arXiv:0904.0969 [hep-ph].

[7] D. Hooper, P. Blasi and P. D. Serpico, JCAP **0**901, 025 (2009) [arXiv:0810.1527 [astro-ph]]; S. Profumo, arXiv:0812.4457 [astro-ph]; D. Malyshev, I. Cholis and J. Gelfand, arXiv:0903.1310 [astro-ph.HE]. H. Yuksel, M. D. Kistler and T. Stanev, Phys. Rev. Lett. **1**03, 051101 (2009) [arXiv:0810.2784 [astro-ph]].

[8] C. R. Chen, F. Takahashi and T. T. Yanagida, Phys. Lett. B **6**71, 71 (2009) [arXiv:0809.0792 [hep-ph]]; P. f. Yin, Q. Yuan, J. Liu, J. Zhang, X. j. Bi and S. h. Zhu, Phys. Rev. D **7**9, 023512 (2009) [arXiv:0811.0176 [hep-ph]]; A. Ibarra and D. Tran, JCAP **0**902, 021 (2009) [arXiv:0811.1555 [hep-ph]].

[9] V. Barger, Y. Gao, W. Y. Keung, D. Marfatia and G. Shaughnessy, Phys. Lett. B **6**78, 283 (2009) [arXiv:0904.2001 [hep-ph]]; X. G. He, arXiv:0908.2908 [hep-ph]; I. Maor, arXiv:astro-ph/0602440; R. Allahverdi, B. Dutta, K. Richardson-McDaniel and Y. Santoso, Phys. Rev. D **7**9, 075005 (2009) [arXiv:0812.2196 [hep-ph]]; G. Huetsi, A. Hektor and M. Raidal, arXiv:0906.4550 [astro-ph.CO]; M. Pospelov and A. Ritz, Phys. Lett. B **6**71, 391 (2009) [arXiv:0810.1502 [hep-ph]]

[10] J. Hisano, S. Matsumoto and M. M. Nojiri, Phys. Rev. Lett. **9**2, 031303 (2004) [arXiv:hep-ph/0307216].

[11] N. Arkani-Hamed, D. P. Finkbeiner, T. R. Slatyer and N. Weiner, Phys. Rev. D **7**9, 015014 (2009) [arXiv:0810.0713 [hep-ph]].

[12] H. Rauch, M. Zawisky, Ch. Stellmach and P. Geltenbort, Phys. Rev. Lett. **83**, 4955 (1999).

[13] K. Griest and D. Seckel, Phys. Rev. D **43**, 3191 (1991).

[14] W. Shepherd, T. M. P. Tait and G. Zaharijas, Phys. Rev.

D **7**9, 055022 (2009) [arXiv:0901.2125 [hep-ph]].

[15] M. Ibe, H. Murayama and T. T. Yanagida, Phys. Rev. D **7**9, 095009 (2009) [arXiv:0812.0072 [hep-ph]].

[16] W. L. Guo and Y. L. Wu, Phys. Rev. D **79**, 055012 (2009) [arXiv:0901.1450 [hep-ph]].

[17] J. F. Beacom, N. F. Bell and G. D. Mack, Phys. Rev. Lett. **99**, 231301 (2007) [arXiv:astro-ph/0608090].

[18] K. A. Milton and I. L. Solovtsov, Mod. Phys. Lett. A **16**, 2213 (2001).

[19] S. J. Brodsky, A. H. Hoang, J. H. Kuhn and T. Teubner, Phys. Lett. B **3**59, 355 (1995) [arXiv:hep-ph/9508274].

[20] Recent work assuming a normalization pre-factor is general include S. Cassel, arXiv:0903.5307 [hep-ph], R. Iengo, arXiv:0903.0317 [hep-ph], and J. Bovy, arXiv:0903.0413 [astro-ph.HE].

[21] E. Guth, C. J. Mullin, Phys. Rev. **83**, 3, 667-668 (1951)

[22] I. Harris and L. Brown, Phys Rev **105**, 1656 (1956). For more recent calculations with great attention to infrared singularities see G. S. Adkins, Y. M. Aksu and M. H. T. Bui, Phys. Rev. A **47**, 2640 (1993).

[23] G. Elwert Ann Phys **3**4, 178 (1939); G. Elwert and E. Haug, Phys Rev **183**, 90 (1969).

[24] E. Guth, Phys Rev **59**, 325 (1941).

[25] R. H Pratt and H. L. Tseng, Phys. Rev. **A**11, 1797 (1975).

[26] L. Hui, Phys. Rev. Lett. **8**6, 3467 (2001) [arXiv:astro-ph/0102349].

[27] E. P. Wigner, Phys. Rev. **73**, 1002 (1949)

[28] Many papers, e.g. Ref. [19] quote results without derivation or approximations. We have not actually found a proof in the literature

[29] Work by Cassel [20] and Iengo [20] seek normalization pre-factors of the form $\sigma_l \to \sigma_l S_l$ from the start. Those assumptions then contradict unitarity, as Cassel noticed.